



## King's Research Portal

DOI:

[10.1111/iej.12679](https://doi.org/10.1111/iej.12679)

*Document Version*

Peer reviewed version

[Link to publication record in King's Research Portal](#)

*Citation for published version (APA):*

Eliyas, S., Briggs, P. F. A., Harris, I. R., Newton, J. T., & Gallagher, J. E. (2016). Development of quality measurement instruments for root canal treatment. *International Endodontic Journal*. DOI: 10.1111/iej.12679

### **Citing this paper**

Please note that where the full-text provided on King's Research Portal is the Author Accepted Manuscript or Post-Print version this may differ from the final Published version. If citing, it is advised that you check and use the publisher's definitive version for pagination, volume/issue, and date of publication details. And where the final published version is provided on the Research Portal, if citing you are again advised to check the publisher's website for any subsequent corrections.

### **General rights**

Copyright and moral rights for the publications made accessible in the Research Portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognize and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the Research Portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the Research Portal

### **Take down policy**

If you believe that this document breaches copyright please contact [librarypure@kcl.ac.uk](mailto:librarypure@kcl.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



## Development of Quality Measurement Instruments for Root Canal Treatment

Journal:	<i>International Endodontic Journal</i>
Manuscript ID	IEJ-16-00049.R3
Manuscript Type:	Original Scientific Article
Keywords:	Outcome of root canal treatment, Quality measures for root canal treatment, Quality of root canal treatment

SCHOLARONE™  
Manuscripts

Review

**Abstract**

Aim: To devise measurement instruments for 'quality' of root canal treatment to assess training and outcome of general dental practitioners working within primary care settings.

Method: Scoring systems relating to quality of root canal treatment were developed using expert consensus and published literature. Domains scored included the Treatment Process, Quality of the Obturation, Clinical Healing, Radiographic Healing and Tooth Complexity. Scoring systems were applied to 10 clinical cases treated by each dentist at the beginning and 10 cases treated at the end of their clinical training and 135 cases treated after completion of training. The dentists recorded the treatment process and clinical healing in clinical logs. Two examiners independently scored the radiographs after undertaking calibration and training. Inter- and intra-examiner reliability of scoring radiographic outcomes was tested.

Results: Instrument created with 4 domains to assess quality (2 process and 2 outcome), and a measure of case complexity (structure). One domain of process (n=240 teeth), one domain of outcome (n=32 teeth) and the complexity (n=215 teeth) were scored using radiographs. The Kappa scores for intra-examiner reliability between 0.22 and 1, whilst inter-examiner reliability ranged between 0.18 and 0.99.

Conclusion: Evidence based scores for assessment of the quality (process and outcome) and complexity (structure) of root canal treatment were devised. They are reliable, provided that clinicians are trained in record keeping and examiners have in depth training and calibration in the use of the instruments.

## Introduction

In root canal treatment, only histological sections allow definitive assessment of healing outcomes (Paula-Silva *et al* 2009, Laux *et al* 2000). Patients measure outcome in relation to the absence of symptoms (Bender *et al* 1996a, 1996b), function and aesthetics (Friedman 2002) and overall quality of life (Dugas *et al* 2002). Insurance companies and dental public health bodies assess survival (presence or absence) of the tooth following root canal treatment (Salehrabi *et al* 2004, Lazarski *et al* 2001, Lumley *et al* 2008, Tickle *et al* 2008, Chen *et al* 2007, Ng *et al* 2010). Clinicians assess the radiographic quality of a root canal filling (quality of obturation) as a surrogate measure of quality of treatment and healing of apical pathology as seen on radiographs as a surrogate end point for outcome. The European Society of Endodontology (ESE) has described the gold standards for root canal treatment as summarised in Table 1 (European Society of Endodontology 2006).

### Table 1

Measurement of anything is the allocation of numbers to the observation being measured. In healthcare these can be theoretical concepts. The instruments used to make the measurement need to have defined indices, which allow the theoretical concept to be allocated numbers that reflect either the presence or absence of the concept or importance of the concept. The quality of the measuring instrument is indicated by how accurately the concept being measured is actually measured (validity) and whether the measurement tool can be used repeatedly to arrive at the same answer if used by any number of trained individuals (reliability) under consistent conditions (Kimberlin & Winterstein 2008).

The development of measurement instruments involves concept development, specifying the dimensions of the concept, selection of indicators and the formation of an index using literature and expert opinion. Using more than one indicator gives stability to the scores and increases their validity; the indicators are then combined to form an index (Kothari 2004).

There are numerous scoring systems and most have measured radiographic healing using the Orstavik classification (Orstavik *et al* 1986) with healing defined slightly differently in each study. Radiographic quality of the obturation has been measured using a variety of subjective definitions. These are simplified and used clinically without information about their reliability. The quality of the root canal treatment provided may be affected by the complexity of the case treated. In order to explore this in future research, a method of easily quantifying complexity is required.

Technical performance in surgery is reflective of both knowledge and judgment used to develop strategies to provide the treatment and the skill involved in implementing those strategies (Darzi & Mackay 2001). These are measured against best practice as determined by the best available knowledge and technology at the time and not ideals unachievable with current knowledge and technology. Quality of care can be classified under 'structure' (facilities, equipment, resources both human and financial, methods of reimbursement), 'process' (what is actually done including the patient seeking care) and 'outcome' (effects of care on health status including the patient's satisfaction with care). Good structure is expected to increase the likelihood of good process, and in turn increase the likelihood of good outcomes (Donabedian 1980, 1966).

In 2009, in line with the Department of Health national policy on Dentists with Enhanced Skills (DES), an innovative collaboration between the London Deanery and what were London Primary Care Trusts (PCTs), formed a training path to provide endodontic patient care in dental practices during a 24 day course over two years and after completion of the training, whilst also providing general dental care (Department of Health & Faculty of GDUK 2004, Department of Health & Faculty of GDUK 2006, Department of Health Primary Care Contracting 2006). An overview of the preliminary research on the scheme in this study has been reported elsewhere (Al-Haboubi *et al* 2014). There is limited evidence in the literature regarding the feasibility of providing such training, the effect of such training on

the dentist's skills and the outcome of root canal treatment within primary care, especially within the United Kingdom. In order to measure these affects, simple and precise measurement instruments are required.

During this study, scoring systems for four domains of quality were developed: quality of clinical treatment process (process); quality of root canal filling as seen radiographically (process); healing as seen clinically (outcome) and healing as seen radiographically (outcome), as well as complexity of teeth treated (structure). This paper describes the development of an objective measure of clinical and radiographic 'quality' for root canal treatment to measure that performed by dental practitioners working in primary care settings. These were closely mapped to that which is carried out in clinical situations daily and those elements that require radiographic assessment were tested for reliability.

## Materials and Methods

**Development of measurement instruments:** All measurement instruments were developed using expert opinion and the currently available literature.

Measuring the complexity of cases: Expert opinion was used to develop a list of characteristics of a tooth, which could be used as a guide to the complexity of treatment. This was then compared with the tooth complexity indices from the American Association of Endodontists, The Royal College of Surgeons of England, Canadian Academy of Endodontics, The Dutch Endodontic Treatment Index and the Endodontic Treatment Classification (Royal College of Surgeons of England 2001, American Association of Endodontists 2005 edited 2010, Falcon *et al* 2001, Canadian Academy of Endodontics 1998, Ree *et al* 2003).

Score for clinical treatment process (measuring the clinical quality of root canal treatment):

Hülsmann *et al* (2005) described the goals of mechanical root canal preparations and achieving these goals during root canal treatment is considered to be performing a high quality root canal filling. The most recent publications from Ng *et al* (2011a, 2011b) relate the findings from a prospective study and outline a list of pre-operative, intra-operative and post-operative factors affecting outcomes of non-surgical root canal treatment (Table 2). The intra operative factors considered important from the literature (Ng *et al* 2011a and European Society of Endodontology 2006) and expert opinion were used to develop a scoring system for the quality of the clinical process of carrying out root canal treatment. Data for the clinical treatment process was ascertained from logbooks maintained by the clinician, maintenance of which was a compulsory part of the course and could be recorded on paper or electronically, following training on how to record the data.

**Table 2**

Score for quality of root canal filling as seen radiographically (measuring the radiographic quality of root canal treatment):

An absence of technical errors, ideal tapered shape of prepared canal with an obturation free of voids extending to within two millimetres of the radiographic apex is a gold standard that is measurable by radiographic means (Friedman 2002, Ng *et al* 2007, Ng *et al* 2008a, Ng *et al* 2008b, de Chevigny *et al* 2008a, de Chevigny *et al* 2008b Farzaneh *et al* 2004). The available literature concerning the current scoring systems and expert opinion was used to develop a list of factors that were thought to denote radiographic quality of obturation in root canal treatments. The course involved teaching on the use of radiographic assessment using film holders as standard to reduce the risk of errors related to film positioning. The quality of the radiograph was assessed using the National Radiation Protection Board guidelines 2001, where score 1 was excellent, score 0 was diagnostically acceptable and a score of -1 was unacceptable (National Radiation

Protection Board 2001). Those radiographs of unacceptable quality (-1) were considered unusable and excluded from further assessment.

Score for healing as seen clinically (measuring clinical outcome): Root canal treated teeth were compared with what is described to be normal, i.e. the lack of pain, swelling, sinus tracts, tenderness to palpation and percussion, tenderness in function and mobility (Friedman 2002, Cohen & Hargreaves 2006). The presence of symptoms, clinical signs and any other negative signs were recorded as part of the logbook maintained by the participants in line with course requirements, following training on clinical record keeping.

Score for the presence of a satisfactory coronal seal: The presence of a satisfactory coronal seal is a measure of process. A dichotomous score for the presence or absence of a satisfactory coronal seal was used (Ng et al 2011a, Ng et al 2008a, Ng et al 2008b, Farzaneh et al 2004, Tickle et al 2008, Salehrabi et al 2004, Aquilino & Caplan 2002).

Score for healing as seen radiographically (measuring radiographic outcome): Radiographic healing was scored using a simple system developed from other scoring systems for healing (Orstavik et al 1986). A similar approach to Ng et al (2011a) was adopted for this study, however, the scoring system was simplified to three possible outcomes: healed, no change and failed (Ng et al 2011a).

**Ethical Approval:** Ethics committee approval (ref no 10/H0718/69) was obtained. Research Governance approval was sought from all seven Primary Care Trusts: Barking and Dagenham PCT (ref no 2298), Ealing and Hounslow PCT, Greenwich PCT (ref no RDGre573), Hammersmith and Fulham PCT, Newham PCT, Kingston PCT and Wandsworth PCT (St George's Healthcare ref: 2010/401K,W) that had a dentist enrolled in the programme and Kings College Hospital as the base trust (ref: KCH11-006).

**Informed Consent:** Patients received information about the study and were invited to participate in the evaluation when they were sent an appointment for treatment with the DES



with the information sheet and the consent form included. These offered opportunities for the patient to discuss the research protocol with either the researchers or the trainee DES. All trainee DESs were made aware of the planned study and much of the data required for the study was collected as a mandatory part of their training recorded in their logbooks. Their consent was formally sought for involvement in the study prior to patient involvement. All trainees worked within primary dental care. Their principal dentist/service manager was asked to provide consent for this study. Consent from patients for anonymised radiographs to be included in the logbook and assessed as part of this study was gained prior to embarking on treatment as part of the consent for being treated by DES during their training period. The inclusion criteria for patients included ability to give informed consent.

**Sample of teeth used:** The teeth for training and calibration included a variety of cases treated by one of the authors (SE) in Year 1 of speciality training intermingled with a random sample of cases treated by the DES during and after their training. The cases treated by the DES during and after training constituted the cases scored for this study using the measurement instruments.

**Assessment of radiographs:** The radiographs collected as part of the logbooks were as per the ESE guidelines (2006). The assessment of the radiographs included plain films photographed on a fluorescent viewing box without magnification and digitised into JPEG format. The digital radiographs were exported from the various digital systems and saved in JPEG form (opinions gathered from two independent radiologists). No measurements were made from the radiographs, therefore saving these files in either RAW or TIFF forms was not requested. The plain films were photographed using a Single Lens Reflex camera (Nikon D90) with the film placed on a bright-light viewing screen in a darkened room. The plain films and digital films were then saved as JPEG images and examined on a single screen (13" MacBook Pro, Apple Inc.) under controlled lighting and viewing conditions.

**Piloting and amendment:** The scoring systems were piloted among experts, general dental practitioners and specialist trainees. The initial scoring system was judged to be overly complex and subjective, and was therefore dichotomised, where possible.

**Training and Calibration:** Two examiners, one internal (SE) and one external (IRH) to the course independently scored all radiographs. Training involved discussion of the scoring system without the involvement of radiographs. Following this, both examiners scored 40 teeth (using radiographs) independently, for complexity, radiographic appearance of obturation and for healing. This number was chosen as a reasonable amount to score to gain an understanding of agreement. The radiographs scored as part of the training and calibration process were used to determine inter and intra examiner reliability. Cohen's Kappa Coefficient (Cohen 1960) scores were calculated, resulting in low values, therefore further training and calibration was carried out. This consisted of jointly examining the previously scored radiographs and discussing the reasons for decision-making in each case where there were differences in scoring. Then a further 30 cases were scored independently by both examiners and inter- as well as intra-examiner reliability testing was carried out. This resulted in improved scores, and once again the cases where examiners scored differently were discussed to enhance their learning. Discussion of cases using radiographs generated a list of notes for the examiners that was used for the actual scoring. Thus each examiner scored the actual cases for this research project independently. Three months following, each examiner re-scored a randomly selected 10% of the radiographs for complexity, quality of root canal filling as seen radiographically and healing as seen radiographically (Figure 1).

### Figure 1

**Randomisation and blinding:** All radiographs were randomised using computer-generated tables to blind the examiners from the clinical treatment process, the clinician and the stage of training of the DES. The examiners were further blinded from the complexity score when

assessing the quality of root canal filling as seen radiographically, and blinded for the quality of the root canal filling as seen radiographically when scoring healing.

**Statistical analysis:** All data for inter- and intra-examiner reliability were initially entered into an Excel (Microsoft Office 2010) spread sheet, verified and analysed using SPSS (IBM Corporation) v22.

**Structure:** The final scoring system for Complexity of Cases is shown in Table 3. The data for the 'number of roots' and the 'length of the root' were gathered from the clinical logbooks. The total complexity score was calculated by addition of the individual domain scores for each tooth. A total score of 3 was considered fairly simple and a score of 18 was considered extremely complex.

**Table 3**

**Process:** The scoring system for clinical treatment process (where the total score could vary from 0=poor, to 5=good) included: use of rubber dam (Y=1, N=0); irrigants (NaOCl + EDTA=2, NaOCl=1, Anything else=0); apex locator (Y=1, N=0); and patency filing (Y=1, N=0). These data were collected from clinical logbooks maintained by the dentists on the course and thus self-reported.

The scoring system for the quality of root canal filling as seen radiographically (where the total could vary from 0=poor to 4=good), included the presence of procedural errors (Y=0, N=1), the root canal filling being within 2mm of rad apex (Y=1, N=0), continuous taper and shape of the preparation (Y=1, N=0) and the presence of voids (Y=0, N=1). The descriptor for procedural errors stated errors as missed canals, access cavity perforations, ledge formation, perforations, strip perforations, canal transportation, zips/hourglass shapes, elbows, canal blockages, separated instruments and foreign objects (Hülsmann *et al* 2005). Continuous taper and shape was defined as being from the apex to the access cavity with

the cross sectional diameter of the canal being narrower at every point apically, the root canal treatment following the shape of the original canal (Schilder 1974).

**Outcome:** The scoring system for healing as seen clinically (where the total could vary from 0=poor to 4=good) included the presence of symptoms (Y=0, N=1), the presence of clinical signs of failure (Y=0, N=1), the presence of any other negative signs (Y=0, N=1) and the presence of a satisfactory coronal restoration (Y=1, N=0). Clinical signs of infection were defined as swelling, sinus, tenderness to palpation and percussion, isolated deep pocket or mobility. Any other negative signs included extraction, fracture and loss of function.

The scoring system for healing as seen radiographically, defined as reduction in size of or no development of an apical area was awarded a score of 2, no change in size of existing apical area was awarded a score of 1 and an increase in size of or development of an apical area was awarded a score of 0. The literature informing these measures can be seen in Table 4.

#### Table 4

### Results

The results are presented for items of structure (case complexity), process (appearance of the root filling as seen radiographically) and outcome (healing as seen radiographically) where examiners scored radiographs. In total, two examiners scored 395 cases independently. The number of cases scored for complexity, obturation and healing are shown in Figure 1.

**Intra-examiner reliability for domains scored using a radiograph:** Intra examiner reliability testing results are shown in Table 5. The scores were good and the agreement

with the final agreed score (T final) improved from the first time of scoring (T1) to the second scoring when 10% were rescored (T2) as shown in Table 6.

**Table 5**

**Table 6**

**Inter-examiner reliability for domains scored using a radiograph:** The scores were initially low, but improved with further training, although it was not maintained (Table 7). This was more notable for Examiner 1.

**Table 7**

The separate domains of quality can be combined to give an overall measurement instrument for quality where 0 is poor quality and 15 is good quality (Table 8).

**Table 8**

**Discussion**

This study contributes to knowledge by assessing the reliability of objective measures for assessing the quality of root canal treatment using periapical radiographs (radiographic appearance of the root canal filling and healing as seen radiographically) and introduces an objective measurement of clinical treatment process of providing root canal treatment (Table 8). The findings suggests that useable and quantifiable quality measures based on current practice can be developed for the outcome of root canal treatment. This is important to have a measure to provide objective feedback to trainees and monitor progress, especially in a new world where measuring quality of outcomes is becoming more important (Darzi 2008), and where training more likely to occur in primary care settings (possibly for specialists as well as dentists with enhanced skills). These informal current practices are also used for

1  
2  
3 triaging referrals for root canal treatment and their reliability will inform the need for regular  
4 training and calibration. Existing scoring systems were not used, as they were considered  
5 complicated and variable without clear reasons for using one scoring system over another.  
6  
7 The factors most often cited and with evidence for impact on outcome were used (Friedman  
8 2002, Ng et al 2007, Ng et al 2008a, Ng et al 2008b, de Chevigny et al 2008a, de Chevigny  
9 et al 2008b, Farzaneh et al 2004).

10  
11  
12  
13  
14  
15  
16 The current study utilised a combination of digital and plain films, much like those that are  
17 referred to specialists for assessment. This will become an important step in triaging as new  
18 patient pathways develop within the NHS (NHS England Introductory Guide for  
19 Commissioning Dental Specialties 2015, NHS England Guide for Commissioning Dental  
20 Specialties – Orthodontics 2015, NHS England Guide for Commissioning Oral Surgery and  
21 Oral Medicine 2015, NHS England Guide for Commissioning Dental Specialties – Special  
22 Care Dentistry 2015, NHS Five Year Forward View 2014). The results highlight the impact of  
23 training and calibration on reliability of scoring plain film radiographs; however, high levels of  
24 agreement were not necessarily maintained over time without repeated training and  
25 calibration. Ideally all radiographs should be viewed on the original screen recommended by  
26 the manufacturer using the software provided with the system and saved in unchangeable  
27 form. It was assumed that the radiographs provided by the course participants were not  
28 altered in any way.

29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44 The inter-examiner reliability scores were high for tooth position and the variance may be as  
45 a result of incorrect entry of data. Treatment type can be deceptive as the presence of  
46 separated instruments can be difficult to determine radiographically and it may not always  
47 possible to determine from a radiograph if the tooth had previously been accessed to  
48 attempt root canal treatment. If in doubt, examiners were advised to present the lowest  
49 score. The Kappa scores for scoring the quality of radiographs were variable ranging from  
50 0.2 to 0.74. Resorption, root curvature, working length and healing received the poorest  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Kappa scores. The improvement seen with further training was not maintained when a much larger number of radiographs were scored. This may reflect a much larger variation in quality of radiographs or difficulty maintaining concentration for lengthy periods of time. Both examiners scored the radiographs in batches of 30-40 to reduce fatigue. Although every effort was made to score the radiographs as soon as possible after training and calibration, due to logistic reasons scoring was completed 4-8 weeks after training and calibration. It was not possible to calculate intra-examiner reliability for healing due to the small number of cases scored.

Other reported scoring of radiographs for the quality of root canal filling, complete independent agreement between all examiners occurred in 32% of cases, with all observers independently arrived at the same periapical diagnosis in 39% of cases and the opinions of all examiners only coinciding in 15% (n=6) of cases (Reit *et al* 1983). In the current study, the agreement between examiners for radiographic scoring ranged from 69.5% to 85.2%; furthermore, inter examiner reliability Kappa scores varied from 0.18 – 0.99 and intra examiner reliability Kappa scores varied from 0.22 – 1. The agreement levels were in excess of 70%. The Kappa scores for measuring healing using a radiograph was low (0.35) as was the agreement level (75%). When intra examiner reliability was measured against the final score (T<sub>Final</sub>) that was agreed for each case (Table 6), there was some improvement in Kappa scores, and agreement, which may reflect the learning that has taken place during discussions of cases to agree a final score.

Arbitrary magnitude guidelines for ideal Kappa scores exist (Petrie & Watson 1999, Landis & Koch 1977, Fleiss 1981). Kappa scores are higher if codes have equal probability of being chosen, if the two observers distribute codes asymmetrically and as the number of codes increases. Therefore no one value of kappa can be regarded as universally acceptable and finding the suitable Kappa values depending on the number of codes, their probability, and observer accuracy is important. For example, given equiprobable codes and observers who

are 85% accurate, the value of Kappa is 0.49 and 0.60, when number of codes is 2 and 3 respectively (Bakeman *et al* 1997). If this is considered the Kappa scores in this current study are acceptable (Tables 5, 6 and 7). It is noteworthy that these Kappa scores may be reflective of the reliability of current clinical practice.

Other similar studies (Dahlström *et al* 2015, Koch *et al* 2015, Dalhstrom *et al* 2011) have assessed the quality of root canal fillings and healing following education in the use of rotary instrumentation. The reported use of treatment techniques were ascertained via questionnaire surveys (Dahlström *et al* 2015, Koch *et al* 2015, Dalhstrom *et al* 2011, Koch *et al* 2009), whereas in the current study the logbook allowed recording of a variety of aspects of root canal treatment in a standardised manner, following training in record keeping. Due to logistic reasons no attempt was made to verify the data in the logbooks with the patient's clinical notes. Therefore there was complete trust in the participants supplying accurate information. In the study by Dahlström *et al* (2015) the reported Kappa scoring was for the appearance of the root canal filling post operatively using a 5 point scale for length, seal and taper of root canal filling. The variability of an ideal tapered shape of a canal may assume less significance in the future with more widespread use of rotary instrumentation. It was not clear if discussion took place or if scoring was independent. The assessment was performed for each root of a tooth. The only procedural error assessed was canal transportation and this was using a dichotomous scale. Dahlström *et al* 2011 reported intra-examiner Kappa scores reaching 0.85 again using the same scale and it was implied that examiners assessed the quality of root fillings together to reach a consensus. These Kappa scores are not comparable with the current study due to the number of points in each scale. Koch *et al* (2015) also assessed the quality of root filling and healing after adoption of rotary instrumentation and single cone obturation in the Public Dental Service in Sweden, using a large sample of teeth before and after training. The inter-examiner Kappa scores for root filling quality at completion of treatment and follow-up were reported as 0.73 and 0.75 for the PAI scores (5 point scale), 0.81 and 0.84 for the density of root canal fillings (dichotomous



scale) and 0.87 and 0.89 for the distance of the root canal filling from the radiographic apex (3 point scale); however it is worth noting that disagreement was present in almost half of the cases assessed and a third examiner was required to reach agreement in 72 cases (Koch *et al* 2015).

A particular problem in the study of general dental practitioners in a busy NHS dental practice is the logistic and financial difficulty in administering a standardised approach to taking radiographs. Although bespoke putty matrices attached to the film holders might be ideal for obtaining reproducible views of teeth to be assessed (to be used each time that particular tooth was to be radiographed), this would be difficult to incorporate into a busy NHS dental practice. Some of the course participants continued to use conventional plain film radiography; others were using digital radiography from the outset, whilst some moved from plain film to digital radiography during the course. Therefore no attempt was made to standardise the radiographic equipment or clinicians with the exception of teaching the use of film holders as standard. Cone Beam Computed Tomography (CBCT) has been shown to have significantly higher sensitivity and specificity compared to plain film and digital radiography; however, as the size of lesion increased the difference in sensitivity and specificity reduced between limited CBCT, indirect intra-oral digital radiography and plain film radiography (Sogur *et al* 2009). There can be an overestimation of root canal treatment success by as much as 30% when using radiography compared to CBCT (Wu *et al* 2009). It is difficult to justify exposing all patients for CBCT examination of root canal filled teeth and it may be some time before CBCT is routine use for the assessment of root canal filled teeth. In the meantime, the potential reliability of current clinical practice in England is reported in this article. It is appreciated that apical periodontitis can be asymptomatic (Lee *et al* 1986), and periapical pathology can exist without apparent radiographic change. Clinical assessment of outcome is based on signs and symptoms, which are subjective, self-reported and very much part of current clinical practice. This sample of teeth scored is

limited as a select group of dental practitioners with an interest in endodontics and desire to develop their skills recruited and supplied the cases assessed within the study.

The development of numerical scoring systems for assessment of Case Complexity is challenging as quantifying complexity is subjective, and aspects of tooth which make treatment complicated are not always cumulative in arriving at a higher complexity score. It is however, important for triaging and pre-treatment assessment. Verification of validity of the complexity instrument is difficult and may not necessarily reflect the true complexity in a meaningful manner as patient factors will play a role that cannot be assessed from radiographs alone. The assessment of the overall complexity including patient factors is beyond the scope of the current study. The proposed scoring system uses data supplied from the clinician regarding length and number of root canals as well as data from the examiners having scored the pre-operative radiograph as is done in most triaging systems and consultation appointments in the NHS to make decisions on complexity. Particular weighting was not given to the domains of resorption or canal obliteration to maintain a dichotomous simple measurement instrument. Therefore the resultant score may be an underestimate of complexity.

In this study, a tooth could score low complexity in most domains and then have a high complexity score for one domain, which would result in the case being categorized as high complexity; however, even with a weighted scoring system the total score could amount to moderate complexity. Therefore it needs to be recognised, that a total quantitative score may not represent true complexity without a qualitative description. This has been illustrated in Table 3 where various minimum, moderate and maximum weighted scores have been allocated to various domains to show the effect on total score.

Previously used scoring systems have allocated numerical weights to the complexity levels, and a sum of the scores has been used to grade complexity (Curtis *et al* 1999, Canadian Academy of Endodontics 1998, Ree *et al* 2003). Assessment of the validity of scoring

instruments for complexity has been attempted, with inconclusive results (Morand 1992, Ree *et al* 2003, Muthukrishnan *et al* 2007). Weighted Kappa for intra-observer agreement was 0.636. Weighted Kappa for inter-observer agreement varied from 0.570 to 0.223. A variety of reasons were highlighted for the 'moderate to poor' reproducibility, including ambiguity and subjectivity (Muthurishnan *et al* 2007).

In this study, the dentist providing the root canal treatment did not always provide the definitive coronal restoration, this was assessed as part of the assessment at follow up. The provision of the definitive coronal restoration is part of Process, whoever in this case was measured at follow-up. It is noted that accurate measurement is difficult however clinical and radiographic assessment is the most appropriate method of assessment (Abbott 2004).

The overriding strength of the study is the fact that data collection and analysis occurred in the 'real world' and mirrors current clinical practice. The measurement instrument developed proved easy to use. Therefore can be used as part of routine data collection in primary and secondary care within the NHS as well as for teaching and training purposes on an international scale, for example this instrument could be used to show that dental graduates are safe starters, for post-qualification training in root canal treatment, as measurement of the abilities and case mix for Dentists with Enhanced Skills, and provides an objective measure of quality and outcome for all clinicians. On a wider scale this study shows the importance of regular training and calibration for all clinicians reporting on radiographs and using radiographs for decision-making or triaging referrals. These mainly dichotomised scores for quality of root canal treatment allow for routine recording of prognostic factors for good outcomes (Ng *et al* 2011a) on a larger scale, which in turn may facilitate reporting of outcomes in NHS dentistry on a larger group of patients and clinicians.

**Conclusion**

An evidence-based measurement tool for the assessment of four dimensions of the quality (process and outcome) of root canal treatment has been devised. The measurement tools using radiographic examination is reliable, provided that the raters have in-depth training and calibration in the use of the tool. These findings highlights a wider problem with individuals assessing radiographs in their day-to-day clinics and making decisions on the complexity of cases to be triaged to different members of staff as well as making decisions on quality and healing. There is therefore a place for regular training and calibration of individuals involved in assessing radiographs and triaging referrals for root canal treatment.

Conflicts of Interest: The authors declare that there are no conflicts on interest.

References

Abbott PV (2004) Assessing restored teeth with pulp and periapical diseases for the presence of cracks, caries and marginal breakdown. *Australian Dental Journal* **49**, 33-39

Al-Haboubi M, Eliyas S, Briggs PFA, Jones E, Rayan RR, Gallagher JE (2014) Dentists with extended skills: the challenge of innovation. *British Dental Journal* **217**, E6. doi: 10.1038/sj.bdj.2014.652. [accessed on 24 April 2016].

American Association of Endodontists (2005) Endodontic Case Difficulty Assessment and Referral.  
[https://www.aae.org/uploadedfiles/dental\\_professionals/endodontic\\_case\\_assessment/2006\\_casedifficultyassessmentformb\\_edited2010.pdf](https://www.aae.org/uploadedfiles/dental_professionals/endodontic_case_assessment/2006_casedifficultyassessmentformb_edited2010.pdf). [accessed on 24 April 2016].

Aquilino SA, Caplan DJ (2002) Relationship between crown placement and the survival of endodontically treated teeth. *Journal of Prosthetic Dentistry* **87**, 256-263

Bakeman R, Quera V, McArthur D, Robinson BF (1997) Detecting sequential patterns and determining their reliability with fallible observers. *Psychological Methods* **2**, 357–370.

Bender IB, Seltzer S, Soltanoff W (1996a) Endodontic Success – a reappraisal of criteria I. *Oral Surgery, Oral Medicine and Oral Pathology* **22**, 780-789.

Bender IB, Seltzer S, Soltanoff W (1996b) Endodontic Success – a reappraisal of criteria II. *Oral Surgery, Oral Medicine and Oral Pathology* **22**, 790-802.

Canadian Academy of Endodontics (1998, reviewed 2006) *Standards of practice*. Edmonton, Canada: Canadian Academy of Endodontics. <http://docplayer.net/14546583-Standards-of-practice.html>. [accessed on 24 April 2016].

Chen SC, Chueh LH, Hsiao CK, Tsai MY, Ho SC, Chiang CP (2007) An epidemiological study of tooth retention after non-surgical endodontic treatment in a large population in Taiwan. *Journal of Endodontics* **33**, 226-229.

Cohen J (1960) A coefficient for agreement for nominal scales. *Education and Psychological Measurement* **20**, 37–46.

Cohen S, Hargreaves KM (2006) *Pathways of the pulp*, 9<sup>th</sup> edn. St. Louis, USA: Mosby Elsevier.

Curtis E K, Simon D C (1999) Endodontic case difficulty assessment: the team approach. *General Dentistry* **47**, 340–344.

Dahlström L, Molander A, Reit C (2011) Introducing nickel-titanium rotary instrumentation in a public dental service: The long-term effect on root filling quality. *Oral Surgery Oral Medicine Oral Pathology Oral Radiology Endodontology* **112**, 814-189.

Dahlström L, Molander A, Reit C (2015) The impact of a continuing education programme on the adoption of nickel-titanium rotary instrumentation and root-filling quality amongst a group of Swedish general dental practitioners. *European Journal of Dental Education* **19**, 23-30.

Darzi A (2008) High Quality of Care for All: NHS Next Stage Review Final Report. Department of Health. <http://www.dh.gov.uk/en/publicationsandstatistics/publications> [accessed on 24 April 2016].

Darzi A, Mackay S (2001). Assessment of surgical competence. *Quality in Health Care* **10**, ii64-ii69.

de Chevigny C, Dao TT, Basrani BR, Marquis V, Farzaneh M, Abitbol S, Friedman S (2008a) Treatment Outcome in Endodontics: The Toronto Study—Phase 4: Initial Treatment. *Journal of Endodontics* **34**, 258–263.

de Chevigny C, Dao TT, Basrani BR, Marquis V, Farzaneh M, Abitbol S, Friedman S (2008b) Treatment Outcome in Endodontics: The Toronto Study— Phases 3 and 4: Orthograde Retreatment. *Journal of Endodontics* **34**, 131–137.

Department of Health, Faculty of GDUK (2004) Implementing a Scheme for Dentists with Special Interests (DwSIs) – London.  
[http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod\\_cons\\_um\\_dh/groups/dh\\_digitalassets/@dh/@en/documents/digitalasset/dh\\_4083120.pdf](http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_cons_um_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4083120.pdf).  
[accessed on 24 April 2016].

Department of Health, Faculty of GDUK (2006a) Guidelines for the appointment of Dentists with Special Interests (DwSIs) in Endodontics: London.  
[http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod\\_cons\\_um\\_dh/groups/dh\\_digitalassets/@dh/@en/documents/digitalasset/dh\\_4133752.pdf](http://webarchive.nationalarchives.gov.uk/20130107105354/http://www.dh.gov.uk/prod_cons_um_dh/groups/dh_digitalassets/@dh/@en/documents/digitalasset/dh_4133752.pdf).  
[accessed on 24 April 2016].

Department of Health. Primary Care Contracting (2006b) Dentists with a Special Interests: a step by step guide to setting up a DwSI service – London. [https://www.pcc-cic.org.uk/sites/default/files/articles/attachments/step\\_by\\_step\\_guidance\\_dwsis.pdf](https://www.pcc-cic.org.uk/sites/default/files/articles/attachments/step_by_step_guidance_dwsis.pdf).  
[accessed on 24 April 2016].

Donabedian A (1980) *The definition of quality and approaches to its management, vol 1: Explorations in quality assessment and monitoring*. Ann Arbor, Michigan, USA: Health Administration Press.

Donabedian A (1966) Evaluating the quality of medical care. *Milbank Q* **44**, 166-203.

Dugas NN, Lawrence HP, Teplitsky P, Friedman S (2002) Quality of life and satisfaction outcomes of endodontic treatment. *Journal of Endodontics* **28**, 819-827.

European Society of Endodontology (2006) Quality Guidelines for endodontic treatment: consensus report of the European Society of Endodontology. *International Endodontic Journal* **39**, 921-930.

Falcon FC, Richardson P, Shaw MJ, Bulman JS, Smith BGN (2001) Developing an index of Restorative Dental Treatment Need. *British Dental Journal*. **190**, 479-486.

Farzaneh M, Abitbol S, Friedman S (2004) Treatment outcomes in endodontics: The Toronto Study. Phases I and II: orthograde re-treatment. *Journal of Endodontics* **30**, 627-633.

Fleiss JL (1981). *Statistical methods for rates and proportions*, 2nd edn. New York, USA: John Wiley.

Friedman S (2002) Prognosis of initial endodontic therapy. *Endodontic Topics* **2**, 59-88.

Hülsmann M, Peters OA, Dummer PMH (2005) Mechanical preparation of root canals: shaping goals, techniques and means. *Endodontic Topics* **10**, 30-76.

Kimberlin CL, Winterstein AG (2008) Validity and reliability of measurement instruments used in research. *American Journal of Health-System Pharmacy* **65**, 2276-2284.

Koch M, Wolf E, Tegelberg A, Petersson K (2015) Effect of education intervention on the quality and long-term outcomes of root canal treatment in general practice. *International Endodontic Journal* **48**, 680-689.

Kothari CR (2004) *Research Methodology: Methods and Techniques*, 2<sup>nd</sup> edn. New Delhi, India: New Age International (P) Ltd.

Landis JR, Koch GG (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**, 159-174.



Laux M, Abbott PV, Pajarola G, Nair PN (2000) Apical inflammatory root resorption: a correlative radiographic and histological assessment. *International Endodontic Journal* **33**, 483–493.

Lazarski MP, Walker WA, Flores CM, Schindler WG, Hargreaves KM (2001) Epidemiological evaluation of the outcomes of non-surgical root canal treatment in a large cohort of insured dental patients. *Journal of Endodontics* **27**, 791-796.

Lee S-J, Messer HH (1986) Radiographic appearance of artificially prepared periapical lesions confined to cancellous bone. *International Endodontic Journal* **19**, 64-72.

Lumley PJ, Lucarotti PSK, Burke FJT (2008) Ten-year outcome of root fillings in the General Dental Services in England and Wales. *International Endodontic Journal* **41**, 577-585.

Marquis VL, Dao TT, DMD, Farzaneh M, Abitbol S, Friedman S (2006) Treatment Outcome in Endodontics: The Toronto Study. Phase III: Initial Treatment. *Journal of Endodontics* **32**, 299–306.

Morand MA (1992) Reliability study of a new evaluation instrument in endodontics. *Journal of Dental Education* **56**, 63 (Abstract # 117).

Muthukrishnan A, Owens J, Bryant S, & Dummer PMH (2007) Evaluation of a system for grading the complexity of root canal treatment. *British Dental Journal* **202**, E26. doi:10.1038/bdj.2007.170 [accessed on 24 April 2016].

National Radiological Protection Board (2001) *Guidance notes for dental practitioners on the safe use of x-ray equipment*. Department of Health Publication. [https://www.gov.uk/government/uploads/system/uploads/attachment\\_data/file/337178/misc\\_pub\\_DentalGuidanceNotes.pdf](https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/337178/misc_pub_DentalGuidanceNotes.pdf) [accessed on 24 April 2016].

NHS Five Year Forward View. (2014). <https://www.england.nhs.uk/wp-content/uploads/2014/10/5yfv-web.pdf>. [accessed on 24 April 2016].

Ng Y-L, Mann V, Gulabivala K (2007) Outcome of primary root canal treatment: a systematic review of the literature – part 1. Effects of study characteristics on probability of success. *International Endodontic Journal* **40**, 921-939.

Ng Y-L, Mann V, Gulabivala K (2008a) Outcome of primary root canal treatment: a systematic review of the literature – part 2. Influence of clinical factors. *International Endodontic Journal* **41**, 6-31.

Ng Y-L, Mann V, Gulabivala K (2008b) Outcome of secondary root canal treatment: a systematic review of the literature. *International Endodontic Journal* **41**, 1026-1046.

Ng Y-L, Mann V, Gulabivala K (2010) Tooth survival following non-surgical root canal treatment: a systematic review of the literature. *International Endodontic Journal* **43**, 171–189.

Ng Y-L, Mann V, Gulabivala K (2011a) A prospective study of the factors affecting outcomes of non-surgical root canal treatment: part 1: periapical health. *International Endodontic Journal* **44**, 583-609.

Ng Y-L, Mann V, Gulabivala K (2011b) A prospective study of the factors affecting outcomes of non-surgical root canal treatment: part 2: tooth survival. *International Endodontic Journal* **44**, 610–625.

NHS England (2015) Introductory Guide for Commissioning Dental Specialties. <https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/09/intro-guide-comms-dent-speci.pdf>. [accessed on 24 April 2016].

NHS England (2015) Guides for commissioning dental specialties – Orthodontics.  
<https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/09/guid-comms-orthodontics.pdf>. [accessed on 24 April 2016].

NHS England (2015) Guide for Commissioning Oral Surgery and Oral Medicine.  
<https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/09/guid-comms-oral.pdf>. [accessed on 24 April 2016].

NHS England (2015) Guides for commissioning dental specialties – Special Care Dentistry.  
<https://www.england.nhs.uk/commissioning/wp-content/uploads/sites/12/2015/09/guid-comms-speci-care-dentistry.pdf>. [accessed on 24 April 2016].

Orstavik D, Kerekes K, Eriksen HM (1986) The periapical index: a scoring system for radiographic assessment of apical periodontitis. *Endodontics and Dental Traumatology* **2**, 20-34.

Paula-Silva FG, Wu MK, Leonardo MR, da Silva LA, Wesselink PR (2009) Accuracy of periapical radiography and cone beam computed tomography scan in diagnosing apical periodontitis using histopathological findings as a gold standard. *Journal of Endodontics* **35**, 1009-12. □

Petrie A, Watson P (1999) *Statistics for veterinary and animal science*, 1st edition; pp. 224–225. Oxford, UK: Blackwell Science.

Real DG, Davidowicz H, Moura-Netto C, Zenkner CLL, Pagliarin CMLP, Barletta FB, de Moura AAM (2011) Accuracy of working length determination using 3 electronic apex locators and direct digital radiography. *Oral Surgery, Oral Medicine, Oral Pathology, Oral Radiology and Endodontology* **111**, e44-e49.

1  
2  
3 Ree MH, Timmerman MF, Wesselink PR (2003) An evaluation of the usefulness of two  
4 endodontic case assessment forms by general dentists. *International Endodontic Journal*  
5 **36**, 545-555.  
6  
7

8  
9  
10 Reit C, Hollender L (1983) Radiographic evaluation of endodontic therapy and the influence  
11 of observer variation. *Scandinavian Journal of Dental Research* **91**, 205-212.  
12  
13

14  
15 Royal College of Surgeons of England (2001) Restorative Dentistry Index of Treatment  
16 Need: Complexity Assessment. [https://www.rcseng.ac.uk/fds/publications-clinical-](https://www.rcseng.ac.uk/fds/publications-clinical-guidelines/clinical_guidelines/documents/complexityassessment.pdf)  
17 [guidelines/clinical\\_guidelines/documents/complexityassessment.pdf](https://www.rcseng.ac.uk/fds/publications-clinical-guidelines/clinical_guidelines/documents/complexityassessment.pdf). [accessed on 24 April  
18 2016].  
19  
20  
21  
22

23  
24 Salehrabi R, Rotstein I (2004) Endodontic treatment outcomes in a large patient population  
25 in the USA: an epidemiological study. *Journal of Endodontics* **30**, 846-850.  
26  
27

28  
29 Schilder H (1974) Cleaning and shaping the root canal. *Dental Clinics of North America* **18**,  
30 269-296.  
31  
32

33  
34 Silveira LFM, Petry FV, Martos J, Neto JBC (2011) In vivo comparison of the accuracy of  
35 two electronic apex locators. *Australian Endodontic Journal* **37**, 70-72.  
36  
37

38  
39 Sogur E, Baksi BG, Grondahl H-G, Lomcali G, Sen BH (2009) Detectability of chemically  
40 induced periapical lesions by limited cone beam computed tomography, intra-oral digital and  
41 conventional film radiography. *Dentomaxillofacial Radiology* **38**, 458-464.  
42  
43  
44  
45

46  
47 Tickle M, Milsom K, Qualtrough A, Blinkhorn F, Aggarwal VR (2008) The failure rate of HNS  
48 funded molar endodontic treatment delivered in general dental practice. *British Dental*  
49 *Journal* 204, E8. doi:10.1038/bdj.2008.133. [accessed on 24 April 2016].  
50  
51  
52  
53  
54  
55  
56  
57  
58  
59  
60

Wu M-K, Shemesh H, Wesselink PR (2009) Limitations of previously published systematic reviews evaluating the outcome of endodontic treatment. *International Endodontic Journal* **42**, 656-666.

For Peer Review

## Figures

Figure 1: The number of teeth scored during this study.

For Peer Review

Tables

Table 1: Summary of the gold standards for root canal treatment, as described by the European Society of Endodontology (European Society of Endodontology, 2006)

Isolation:	By the use of rubber dam
Determining the working length	Use electronic and radiographic methods to determine working length (should be as close to the apical constriction as possible – i.e. between 0.5 and 2mm of the radiographic apex). It may be necessary to take more than one working length radiograph.
Preparation of the root canal system	The prepared canal should include the original canal, the apical constriction should be maintained, the canal should end in an apical narrowing, the canal should be tapered from crown to apex
Irrigation	The irrigant solution should preferably have disinfectant and organic debris dissolving properties, should be delivered in copious amounts as far up the canal as possible without risking extrusion beyond the foramen, and may be delivered by ultrasonic or sonic systems
Obturation of the root canal system	The quality of the filling must be checked with a radiograph which should show the root apex and preferably 2-3mm of the periapical region. The filled canal should be completely filled unless a post space is required and contain the original canal. No space should be seen between the canal filling and the canal walls. There should be no canal space visible beyond the end point of the root canal filling.
Assessment of outcome of root canal treatment	<p>Should be assessed at least after 1 year and subsequently as required.</p> <p><u>Favourable outcome:</u> absence of pain, swelling and other symptoms, no sinus tract, no loss of function and radiological evidence of a normal periodontal ligament around the root.</p> <p><u>Uncertain outcome:</u> periapical lesion remains the same size or has only reduced in size. In this situation it is recommended that the lesion is further monitored for a minimum period of 4 years. If the lesion persists, the tooth may be associated with post-treatment disease.</p> <p><u>Unfavourable outcome:</u> tooth is associated with signs and symptoms of infection, a radiologically visible lesion has appeared subsequent to treatment or a pre-existing lesion has increased in size, the lesion has remained the same size or only diminished in size during the 4 year assessment period, or continuing root resorption is present.</p> <p><i>Exception: the presence of scar tissue – an extensive radiological lesion may heal but leave a locally visible, irregularly mineralised area. This tooth should continue to be assessed.</i></p>

Table 2: Summary of factors affecting outcome of root canal treatment (Ng *et al* 2011a).

Study	Success rates	Conditions found to improve periapical healing
Success rate of primary root canal treatment (Ng <i>et al</i> 2011a)	83% (95% CI: 81%, 85%)	1. The pre-operative absence of periapical lesion 2. In presence of periapical lesion, the smaller its size 3. The absence of a pre-operative sinus tract 4. Achievement of patency at the canal terminus 5. Extension of canal cleaning as close as possible to its apical terminus
Success rate of secondary root canal treatment (Ng <i>et al</i> 2011a)	80% (95% CI: 78%, 82%)	6. The use of EDTA solution as a penultimate wash followed by a final rinse of NaOCl in secondary root treatment cases 7. Abstaining from using 2%CHX as an adjunct irrigant to NaOCl solution 8. Absence of tooth/root perforation 9. Absence of inter appointment flare up (pain or swelling) 10. Absence of root canal filling extrusion 11. Presence of satisfactory coronal restoration



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

Table 3: Scoring system for the Complexity of Teeth Treated. The first row of results represents the current scoring system. Rows 2 and 3 of results illustrate the effect on the total score if weighting is added to specific domains. Row 2 uses a minimal complexity tooth as an example and row 3 used a high complexity tooth.

Code	Quality of pre op radiograph	No of roots (One = 1 Two = 2 Tree = 3 Four = 4 Five+ = 5	Position in mouth (Up Ant = 1 Low Pos = 2 Low Ant = 3 Up Pos = 4)	Type of Tx Denovo Tx = 1 ReTx = 2 Post removal = 3 Open apex = 4 Pre-op procedural error = 5	Resorption* (Y=1, N=0)	Root curvature* >35' (Y=1, N=0)	Root length >25mm^ (Y=1, N=0)	Canal not visible in any part of canal* (Y=1, N=0)	Total (3 -18)
1		1	1	1	0	0	0	0	3
2		1	1	1	5	0	0	0	8
3		5	4	5	5	5	2	5	31

Key: Pre-op = pre-operative, Up Ant = upper anterior, Low Pos = lower posterior, Low Ant = lower anterior, Up Pos = upper posterior, Tx = Treatment, Denovo Tx = Primary root canal treatment, ReTx = Secondary root canal treatment,

\* Weighted score of 5 for Y=1 as these domains are considered high complexity

^ Weighted score of 2 for Y=1 as this domain is considered moderate complexity

Table 4: The scoring systems developed for quality of clinical treatment process, quality of root canal filling as seen radiographically and healing as seen radiographically.

Scoring System in this study	Gold standard	Supporting Literature
<b>Clinical treatment process</b> (Recorded by clinician, blinded to what is being assessed)	Rubber Dam used Yes=1, No=0 Irrigants NaOCl + EDTA = 2 NaOCl=1 Anything else=0 Apex Locator used Yes=1, No=0 Patency filing Yes=1, No=0	Rubber dam is used NaOCl with penultimate wash with EDTA and final wash with NaOCl Use apex locator to determine apical terminus Gain and maintain patency during treatment European Society of Endodontology Guidelines, 2006 Ng <i>et al</i> 2011a – 0.2% CHX reduces odds of success by 53%. EDTA has no effect on primary RCT but increases odds of success in secondary RCT by 2x European Society of Endodontology Guidelines, 2006 Real <i>et al</i> 2011 - accuracy of finding apical terminus with apex locators 92% vs digital radiographs 65% Silveira <i>et al</i> 2011 – accuracy of finding apical terminus with apex locators 82-92% Ng <i>et al</i> 2011a – if patency gained 2x as likely to have success
<b>Quality of root canal filling as seen radiographically</b> (Examiner-assessed, randomised and clinician and stage of training)	Procedural errors Yes=0, No=1 Within 2mm of rad apex Yes=1, No=0	No procedural errors: missed canals, access cavity perforations, ledge formation, perforations, strip perforations, canal transportation, zips/hourglass shapes, elbows, canal blockages, separated instruments and foreign objects (Hülsmann <i>et al</i> 2005). Obturation must be in the canal within 2mm of the radiographic apex Ng <i>et al</i> 2011a – pre-operative root perforation reduces odds of success by 56% Marquis <i>et al</i> 2006 – healing better if no intra-operative complications (OR=2) de Chevigny <i>et al</i> 2008a – mid treatment complications reduce rate of healing by 15% in primary RCT de Chevigny <i>et al</i> 2008b – pre-operative perforation reduces outcome Farzaneh <i>et al</i> 2004 – pre-operative perforation reduces outcome by OR of 27 in secondary RCT Farzaneh <i>et al</i> 2004 – root canal filling 0-2mm from radiographic apex is better than long root canal filling especially if pre-operative apical area present European Society of Endodontology Guidelines, 2006

		Ng <i>et al</i> 2008a – For primary RCT: root canal filling length affects outcome especially if an apical area already exists. Flush root canal filling > short root canal filling > long root canal filling, if no apical area. Lowest success rate if apical area + short or long Ng <i>et al</i> 2008b – For secondary RCT: short root canal filling > flush root canal filling > long root canal filling (worse if apical area also present) Ng <i>et al</i> 2011a – Odds of success reduced by 12% for every 1 mm short of the radiographic apex. Odds of success reduced by 62% if the root canal filling was long	
Continuous taper and shape: Yes=1, No=0		From the apex to the access cavity, with the cross sectional diameter of the canal being narrower at every point apically, the root canal filling following the shape of the original canal (Schilder 1974)	European Society of Endodontology Guidelines, 2006 Schilder 1974
Voids: Yes=0, No=1		No voids in the obturation	European Society of Endodontology Guidelines, 2006
Healing as seen radiographically (Examiner-assessed, randomised and blinded to clinician)	Reduced or no development of an apical area = 2 No change in size of existing apical area = 1 Increased or development of an apical area = 0	Reduction or no development of an apical area	European Society of Endodontology Guidelines, 2006 Orstavik <i>et al</i> 1986
Healing as seen clinically (Recorded by clinician, blinded to what is being assessed)	Symptoms: Yes=0, No=1 Clinical signs: Yes=0, No=1 Any other negative signs: Yes=0, No=1	Elimination of all clinical signs and symptoms of infection	Friedman 2002, Cohen & Hargreaves 2006
Coronal seal as	Satisfactory coronal restoration:	Provision of a satisfactory	Ng <i>et al</i> 2011a, Ng <i>et al</i> 2008a, Ng <i>et al</i> 2008b, Farzaneh <i>et al</i>

<b>seen clinically and radiographically</b> (recorded by clinician, blinded to what is being assessed)	Yes=1, No=0	coronal seal	2004, Tickle <i>et al</i> 2008, Salehrabi <i>et al</i> 2004, Aquilino & Caplan 2002
---	-------------	--------------	---

Table 5: Intra Examiner reliability for scoring using radiographs for clinical cases

Intra examiner reliability		All Clinical Cases			
		Examiner 1		Examiner 2	
		Kappa	%	Kappa	%
Obturation (n=24 teeth)	Procedural errors	0.51	87	0.33	88
	Working length	0.82	91	0.05	63
	Continuous taper	0.6	83	0.07	54
	Voids	0.72	87	0.74	88
Complexity (n=21 teeth)	Resorption	0.38	79	0.35	84
	Root curvature	0.22	75	0.5	84
	Sclerosis	0.58	80	0.87	94
	Position	1	100	1	100
	Type of tx	0.91	95	1	100
Healing (n=3 teeth)		*	100	*	100

\* Not able to be calculated due to the lack of significantly different scores

Table 6: Intra Examiner reliability of examiners for scoring using radiographs when compared to the agreed final score for clinical cases scored for the study after training and calibration

Intra Examiner Reliability		T1* vs T2^				T1* vs T final^#				T2^ vs T final^#			
		Examiner 1		Examiner 2		Examiner 1		Examiner 2		Examiner 1		Examiner 2	
		K	%	K	%	K	%	K	%	K	%	K	%
Obturation (n=24 teeth)	Procedural errors	0.51	87	0.33	88	0.51	88	0.50	87	0.86	96	0.25	83
	Working length	0.82	91	0.05	63	0.82	91	0.26	70	0.82	91	0.35	70
	Continuous taper	0.60	83	0.07	54	0.6	83	0.65	83	0.82	91	0.20	61
	Voids	0.72	87	0.74	88	0.82	91	1	100	0.91	96	0.82	91
Complexity (n=21 teeth)	Resorption	0.38	79	0.35	84	0.6	90	0.46	90	0.69	90	0.83	95
	Root curvature	0.22	75	0.50	84	0.27	80	1	100	0.88	95	0.50	95
	Sclerosis	0.58	80	0.87	94	0.55	80	0.73	89	1	100	0.82	95
	Position	1	100	1	100	1	100	1	100	1	100	1	100
	Type of treatment	0.91	95	1	100	0.83	91	0.91	95	0.91	95	0.89	95

\*T1 = the first set of scores by each examiner

^T2 = the second set of scores by each examiner, performed 3 months after T1

#T Final = the final scores agreed for the study

Table 7: Inter Examiner reliability for scoring using radiographs

Inter examiner reliability		After training (n=40 teeth)		After further training + calibration (n=30 teeth)		All cases for study (n=240 teeth)	
		Kappa	%	Kappa	%	Kappa	%
Obturation	Procedural errors	0.56	84	0.44	86	0.37	85
	Working length	0.37	68	0.31	68	0.29	71
	Continuous taper	0.35	72	0.66	86	0.38	70
	Voids	0.44	74	0.13	79	0.54	78
						n=215 teeth	
Complexity	Resorption	0.39	85	0.57	83	0.26	86
	Root curvature	0	95	-0.05	83	0.18	87
	Sclerosis	0.54	78	0.65	83	0.59	79
	Position	1	100	1	100	0.99	99
	Type of tx	0.83	89	0.64	78	0.85	91
						n=32 teeth	
Healing		0.19	72	0.51	81	0.35	75

Table 8: The criteria used for the measurement of Process and Outcome as described by Donabedian (1980, 1966)

<b>Measurement of Process</b>	Clinical Treatment Process score	Rubber Dam used (Y=1, N=0)
		Irrigants (NaOCl + EDTA = 2, NaOCl=1, Anything else=0)
		AL used (Y=1, N=0)
		Patency filing (Y=1, N=0)
<b>Measurement of Process</b>	Quality of root canal filling as seen radiographically	Procedural errors (Y=0, N=1)
		Within 2mm of rad apex inside the root canal (Y=1, N=0)
		Continuous taper and shape (Y=1, N=0)
		Voids (Y=0, N=1)
<b>Measurement of Outcome</b>	Healing as seen radiographically	12 month Healing - Apical area (Reduced or no development of an apical area =2, no change in size of existing apical area =1, Increased or development of an apical area =0)
<b>Measurement of Outcome</b>	Healing as seen clinically	Symptoms (Y=0, N=1)
		Clinical signs (Y=0, N=1)
		Any other negative signs (Y=0, N=1)
<b>Measurement of Process</b>	Quality of the coronal seal as seen clinically and radiographically	Satisfactory coronal restoration (Y=1, N=0)
		Total quality score (0=poor, 15=good)



1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41  
42  
43  
44  
45  
46  
47  
48  
49

For Peer Review

Figure 1: The number of teeth scored during this study.

